LA-UR-22-20888

Approved for public release; distribution is unlimited.

Title:	P. putida Learn Experiments – TrRNA library – LANL & ANL
Author(s):	Davison, Steffi Angela Kern, Theresa Lynn Dale, Taraka T.
Intended for:	Meeting with collaborators
Issued:	2022-02-01

Agile BioFoundry

U.S. DEPARTMENT OF ENERGY | Energy Efficiency & Renewable Energy

BIOENERGY TECHNOLOGIES OFFICE

P. putida Learn Experiments – TrRNA library – LANL & ANL

Steffi Davison, Theresa Kern, Naresh Panday, Claire Sanders, Bert Huttanus, Gyorgy Babnigg, Phil Laible, Rose Wilton, Debbie Hanson, Ramesh Jha, Scott Hennelly, Taraka Dale

31 January, 2022

LA-UR-20-30268

Background and context

ABF libraries in past

- *Cis*-repressor library (LANL, Scott Hennelly and Naresh Pandey)
- Biosensor libraries (LANL Ramesh)
- RB-Tn-Seq (ORNL, Adam Guss)
  - Adam's TnSeq library was sorted a couple of times
  - Conceptually this should work, but there were technical issues that we didn't completely work through
  - low [DNA] led to insufficient sequence depth

In progress ABF Libraries + point of contact

- IMRE library (LANL, Theresa Kern)
- RB-Tn-Seq library (NREL, Caroline Amendola)
- CRISPRi library (Oakridge, Carrie Eckert)

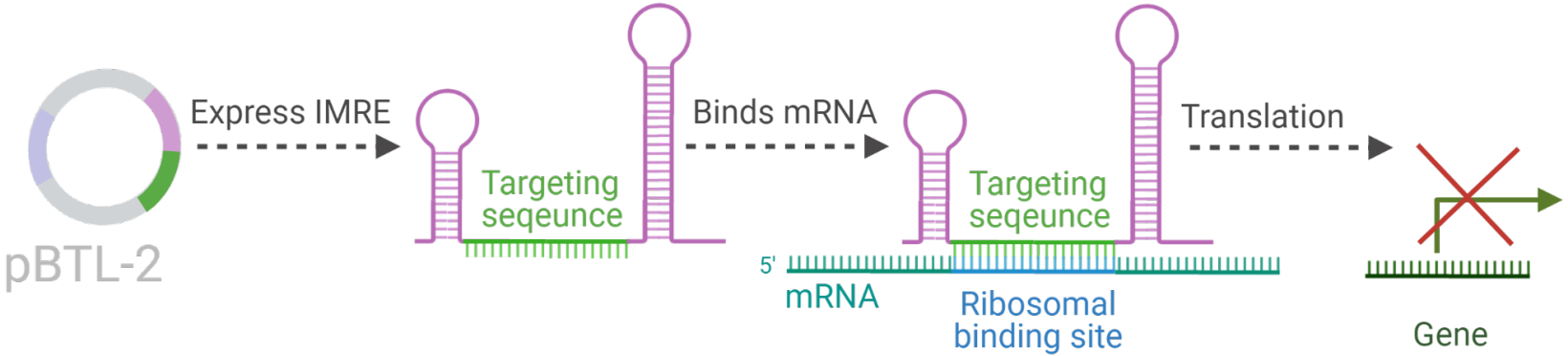
How does IMRE library differ from other *putida* work?

- How is this different from the ongoing work?
  - No ALE or random mutagenesis
  - So far, sensor/sorting work has focused only on 'top performers'
  - Here, we will also collect 'poor and intermediate' performance' data: important for LEARN to bound what doesn't work
  - Ideally this work will also give strains that increase TRY, but the main goal is to give **LEARN a richer dataset to improve Learn outcomes for increasing TRY**

- How is this different from using a transposon or other KO library?
  - Conceptually the same! But practically it's different
  - Here we will use a **plasmid-based approach**, which is easier for DNA isolation & sequencing, AND we've already done it before
  - Also interested in seeing how **knockdowns perform versus knockouts**

U.S. DEPARTMENT OF ENERGY | Energy Efficiency & Renewable Energy
BIOENERGY TECHNOLOGIES OFFICE
Agile BioFoundry

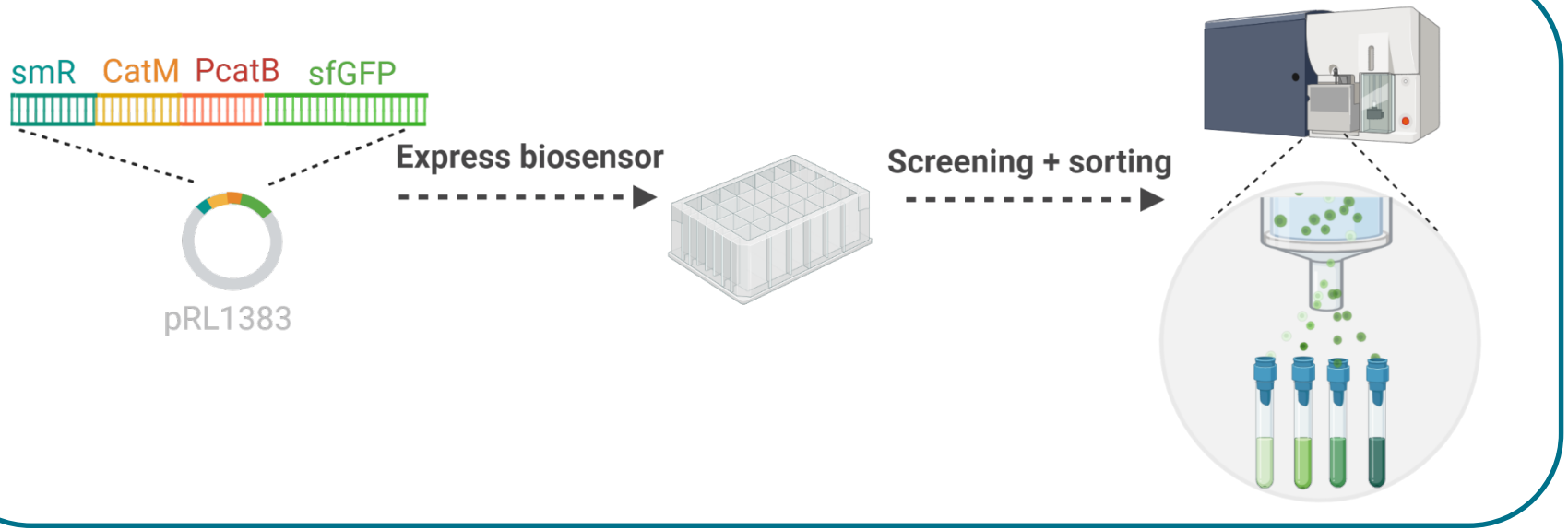
IMRE Riboregulator RNA structures (Naresh Pandey and Scott Hennelly)



- TrRNA, a Twist library of all *P. putida* genes assembled
- Knockdown all ~5500 genes in *P. putida*

Riboregulators controls RBS accessibility, in turn, allows tunable protein expression control of a target mRNA '*in trans*'

Screening with biosensor work (Ramesh Jha and Taraka Dale)



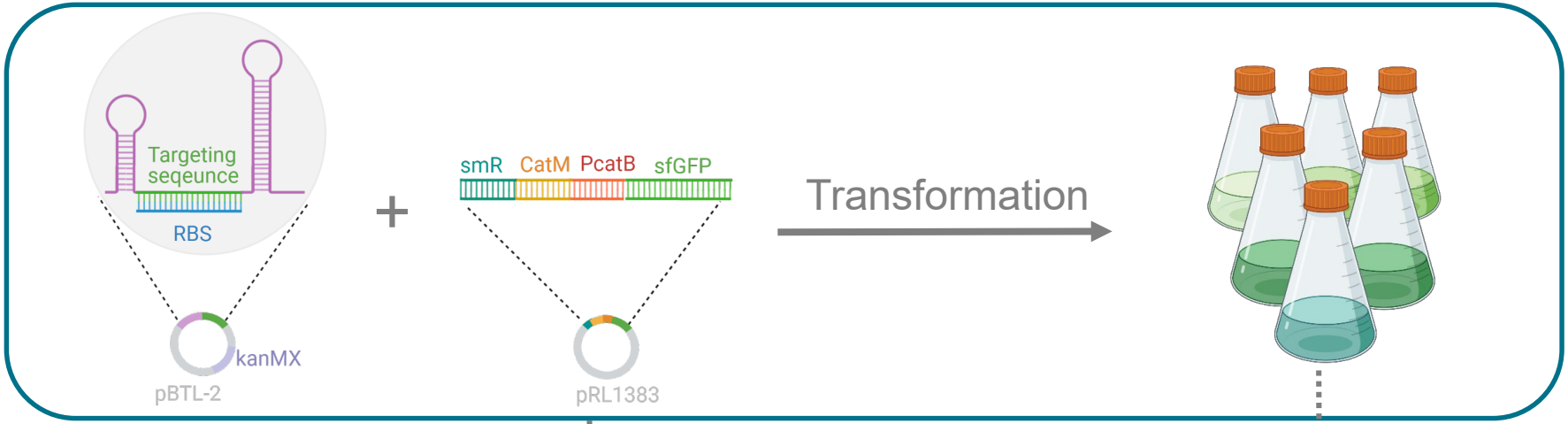
Use fluorescent biosensors & flow cytometry to rapidly screen and isolate populations of strains with varied phenotypes

Hypothesis: Adding “negative”/”poor performance” data to the Learn models will better inform what makes a good performer, leading to improved predictions for increased TRY.

The Learn team has been saying this for a while, but experimentally the *Putida* team has thus far only focused on generating and characterizing and better strains, we have not placed resources on looking at poor performers

Goal: Provide the Learn team with a richer dataset, resulting from isolating and characterizing *P. putida* strains with a range of phenotypes for muconate production, including poor performers.

General experimental workflow



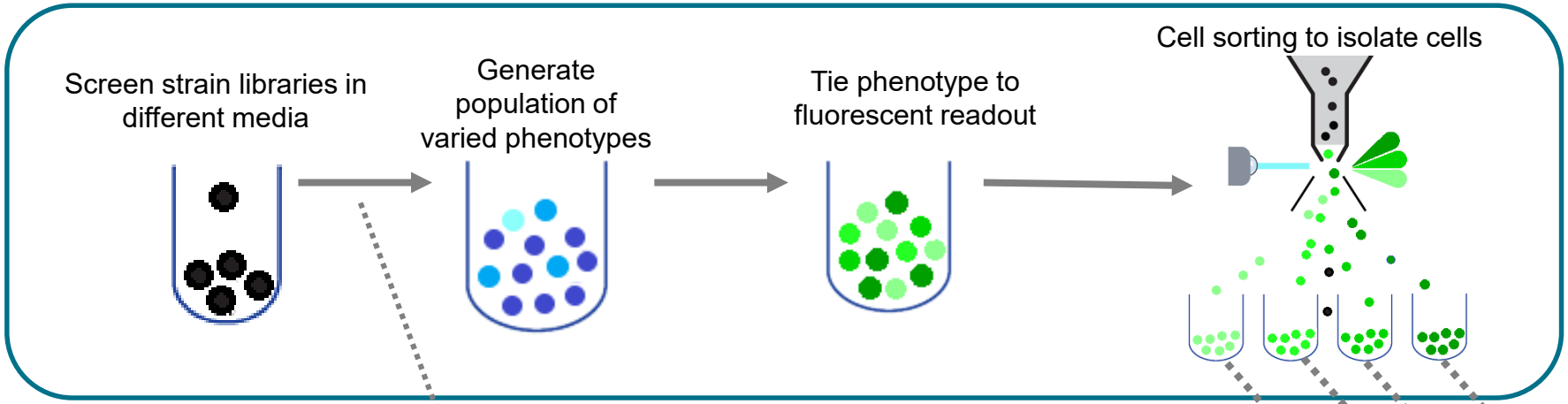
TrRNA knockdown library (pBTL-2+kanMX)

Assay will use CatM C2.9-sfGFP biosensor (pRL1383+smR)

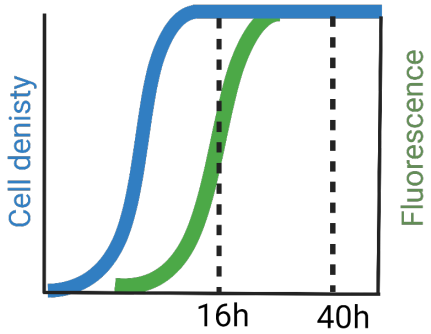
Muconate-producing *P. putida* strains

1. NP162
2. SD3
3. GB271
4. LC238
5. LC236

General experimental workflow plan

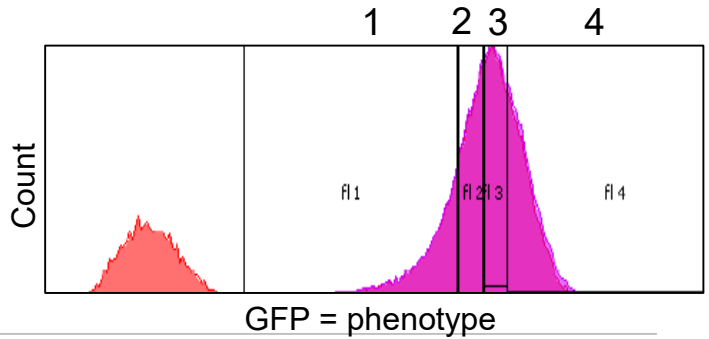


Lag in GFP signal
vs
Drop out rate



Define populations to be collected by "gating"

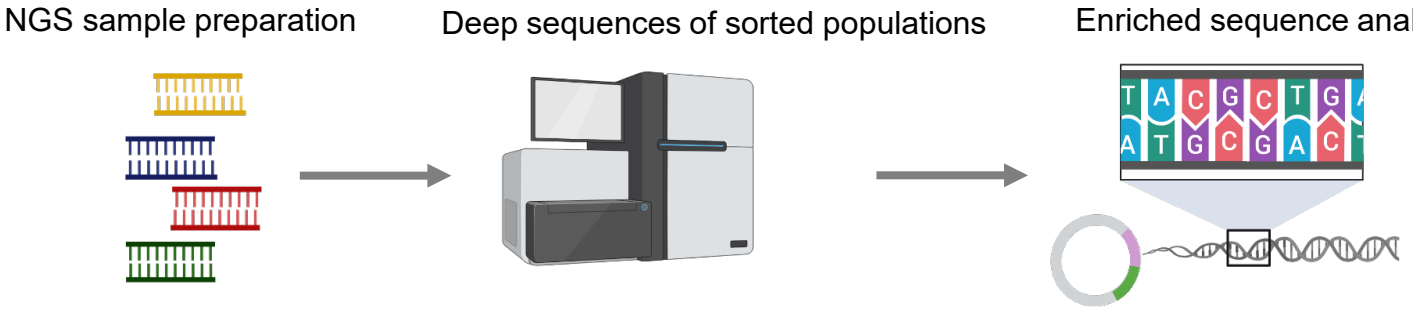
= 2 mil. cells per population



U.S. DEPARTMENT OF ENERGY | Energy Efficiency & Renewable Energy
BIOENERGY TECHNOLOGIES OFFICE

Agile BioFoundry

General experimental workflow plan

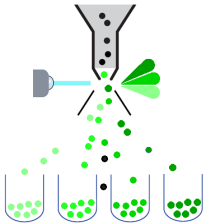


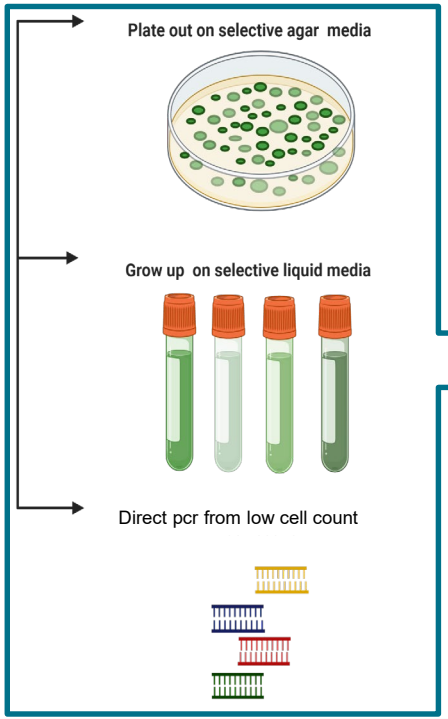
- Plasmid sequencing of each sorted population to look for enriched sequences that can be correlated to each phenotype
- Identify genes of interested from enriched pools
- Finally, validate the genes in ‘clean’ background

U.S. DEPARTMENT OF ENERGY | Energy Efficiency & Renewable Energy
BIOENERGY TECHNOLOGIES OFFICE
Agile BioFoundry

Outgrowth and PCR Bias Investigation

1 step barcode/adapter pcr





E. coli plasmid library

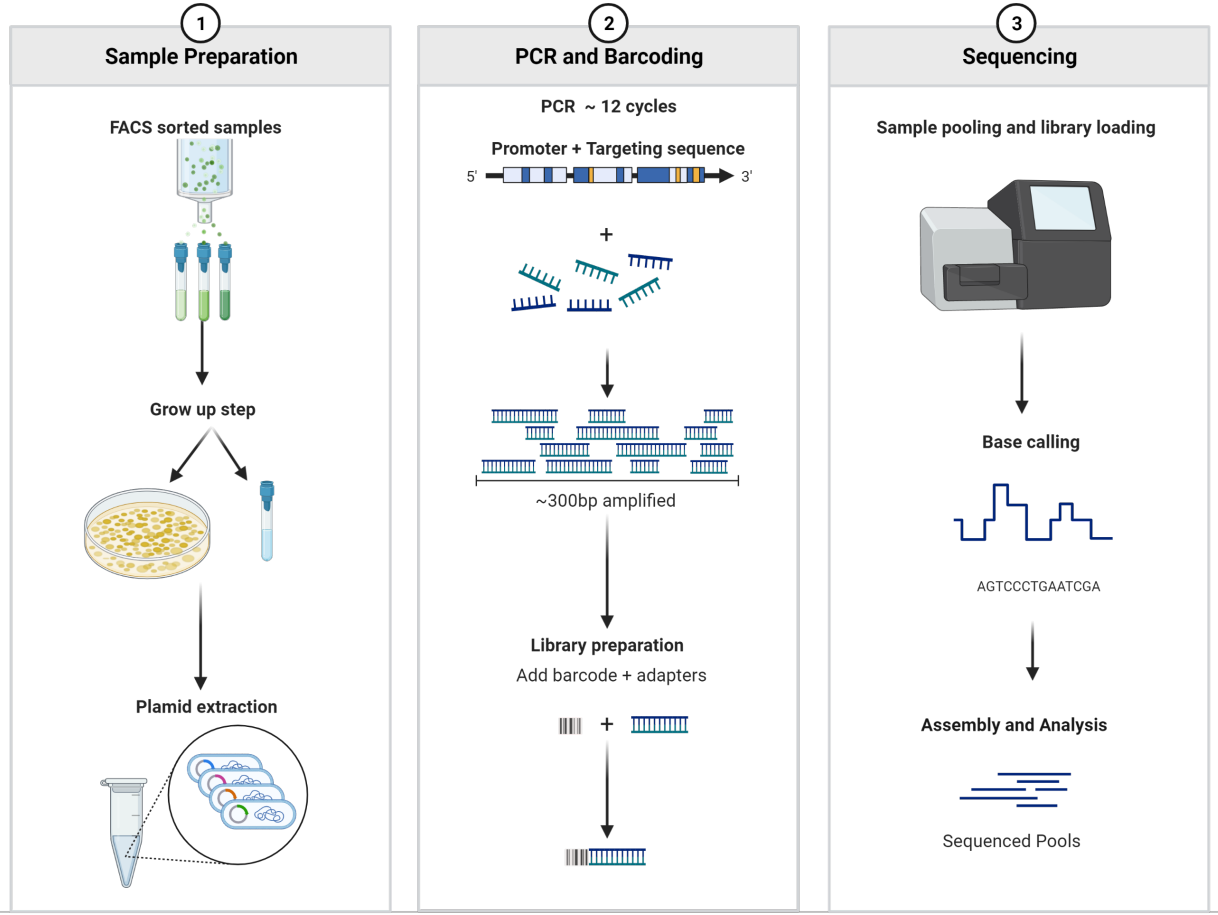
0, 5,10,15,20,25 cycles

Minimum # of cycles attempted. Increased due to low yields

- Gel extraction cleanup
- Qubit /Agilent Bioanalyzer
- Normalize and pool

LBL MiSeq

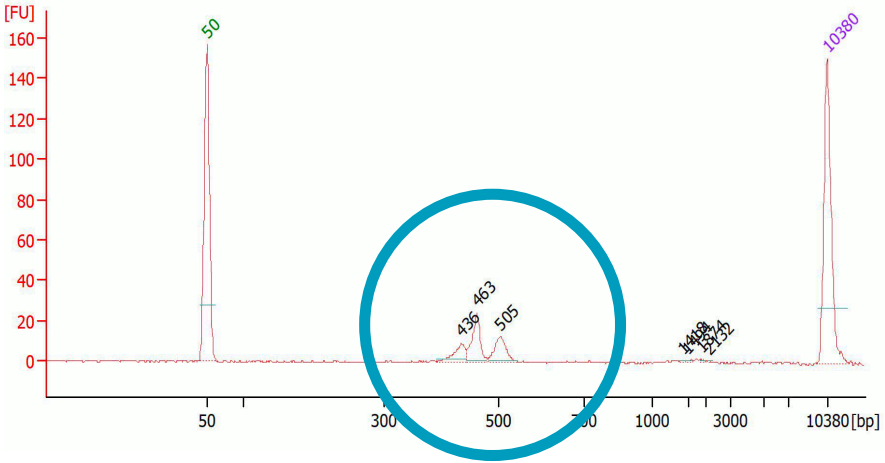
Library Deep Sequencing using Illumina MiSeq Technologies

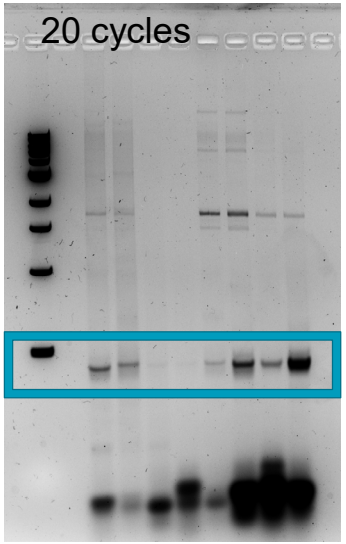


U.S. DEPARTMENT OF ENERGY | Energy Efficiency & Renewable Energy
BIOENERGY TECHNOLOGIES OFFICE

Challenges with barcoding, library pool prep

- Plasmid 1 Step PCR: Meant to simplify and expedite process
  - Long primers : Fwd_illumina_ppIMRE_A501 AATGATACGGCGACCACCGAGATCTACACTGAACCTT ACACTCTTTCCCTACACGACGCTCTTCCGATCT CGC GGC ACA CTC TAG AGT TT
  - **Presence of primer dimers/nonspecific bands and low target amplicon yields:** time consuming and tedious library prep.
    - Multiple PCRs with high cycle number
    - Gel band excision cleanup

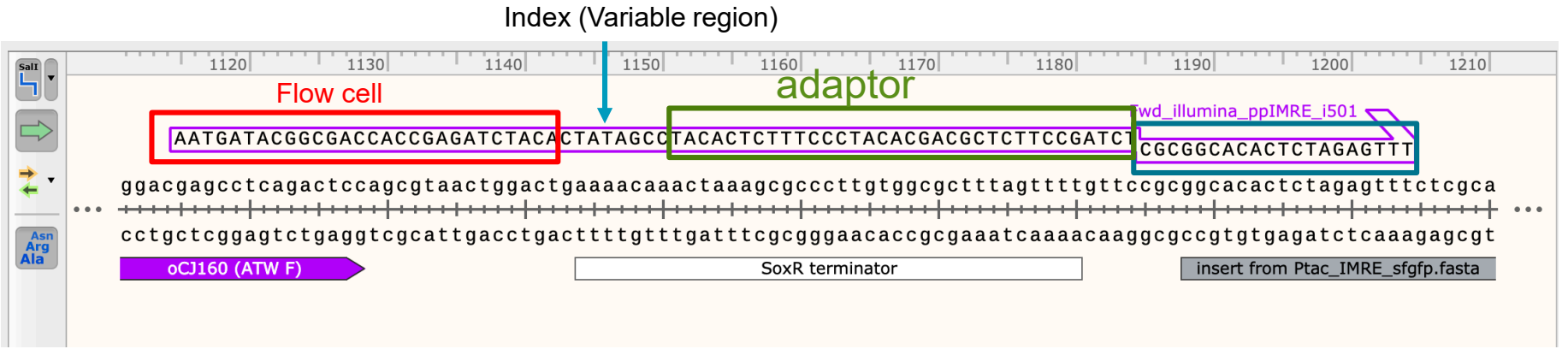




U.S. DEPARTMENT OF ENERGY | Energy Efficiency & Renewable Energy
BIOENERGY TECHNOLOGIES OFFICE

Troubleshooting/protocol discussion with LBL

- Preferential formation of dimers/concatamers
- Low complexity of amplicons
  - Difficult for MiSeq analysis



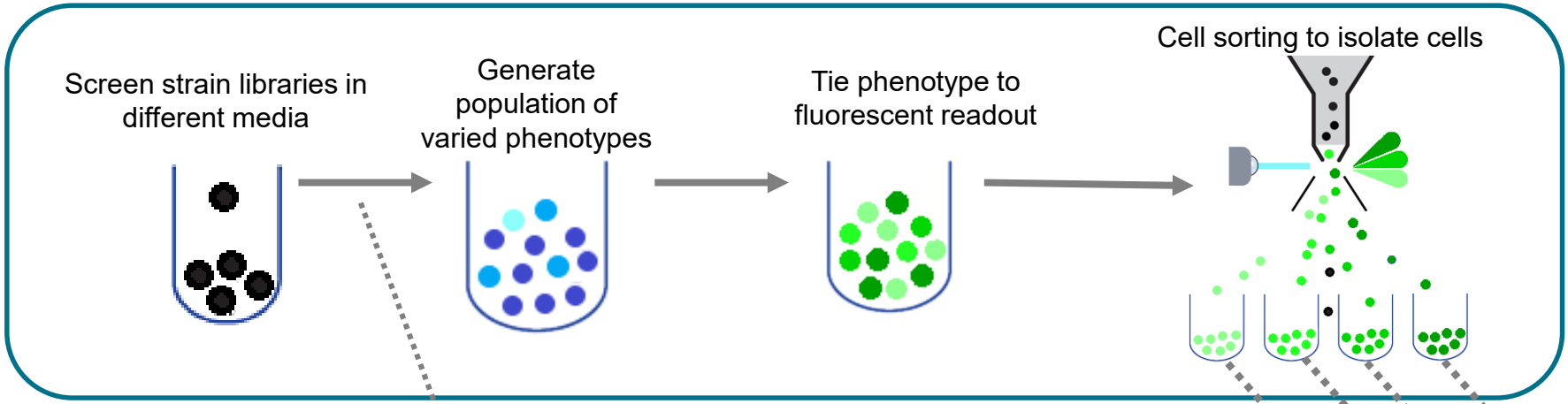
U.S. DEPARTMENT OF ENERGY | Energy Efficiency & Renewable Energy
BIOENERGY TECHNOLOGIES OFFICE

Next steps for IMRE library

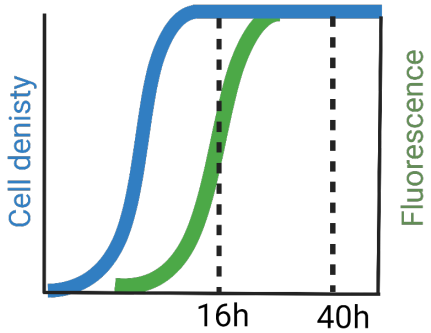
- Repeat Library growth/sorting/ pcr bias with “more sensitive” muconate sensor. -This should reduce GFP signal lag time
- Improved Miseq library pool prep with guidance from LBL:
  - Several options being explored:
    - 1 step pcr with altered long primers: removal of extra “G: residue in P7 adapter - i7 index - Read 2 primer region. (Can inhibit template binding)

      5’ CAAGCAGAAGACGGCATACGAGAT- index - G TGACTGG AGTTCAGACG TGTGCTCTTCCGATCT---AAC AGA TAA AAC GAA AGG CCC 3’
    - 2 step PCR with short primers/limited cycle numbers
    - 10 cycle of PCR + Illumina adapter
    - Or instead of 2nd pcr, adapter ligation starting with 100 ng from 1st pcr
    - Improved Thermocycling protocols
    - Magnetic bead cleanup

General experimental workflow plan

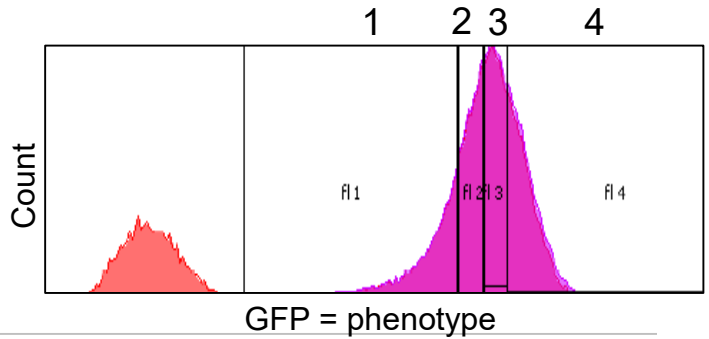


Lag in GFP signal
vs
Drop out rate



Define populations to be collected by "gating"

= 2 mil. cells per population



U.S. DEPARTMENT OF ENERGY | Energy Efficiency & Renewable Energy
BIOENERGY TECHNOLOGIES OFFICE

RB_Tn-Seq library (NREL)

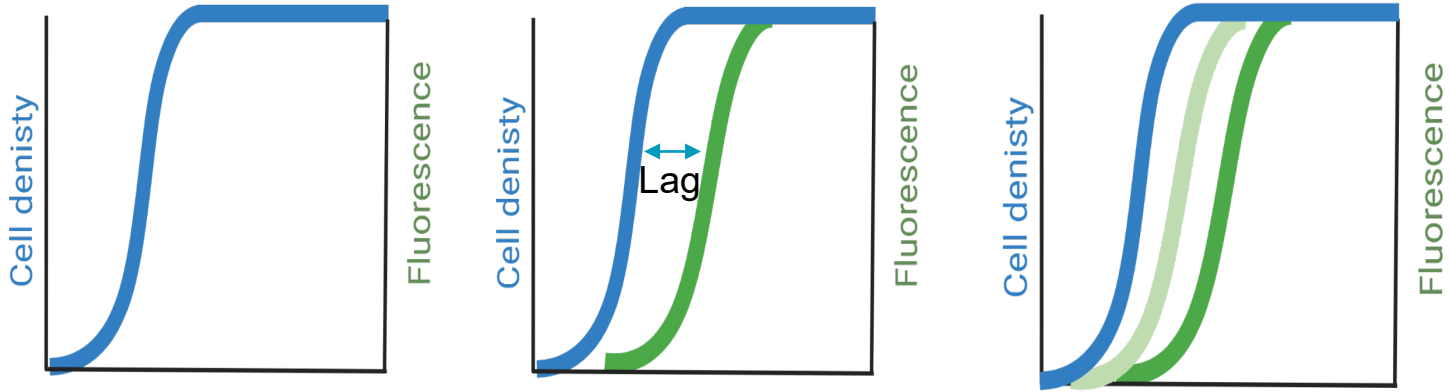
Muconate biosensors:

- **CA089 and CA090** - CJ781 (muconate from aromatics)
  - M9 minimal media with 20 mM p-coumarate and 10 mM glucose
- **CA091 and CA092** – KH083 (muconate from sugars)
  - M9 minimal media with 30 mM glucose and 15 mM xylose

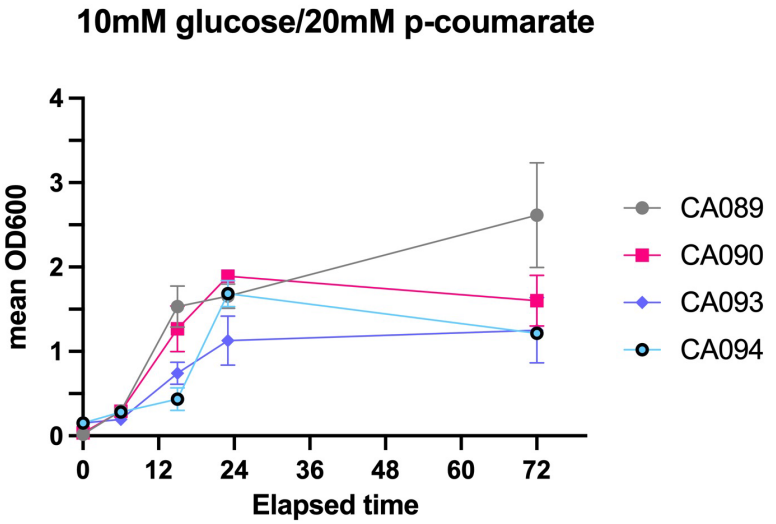
BKA biosensors:

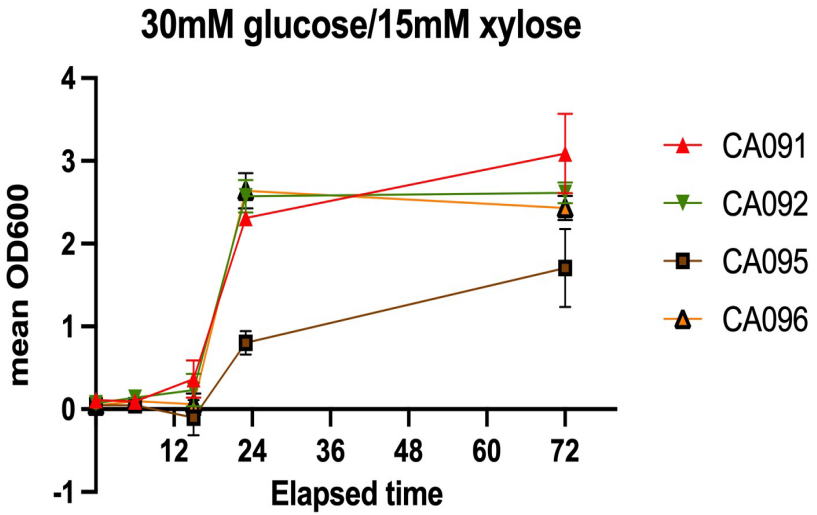
- **CA093 and CA094** – AW311 (BKA from aromatics)
  - M9 minimal media with 20 mM p-coumarate and 10 mM glucose
- **CA095 and CA096** – GR038 (BKA from sugars)
  - M9 minimal media with 30 mM glucose and 15 mM xylose

Biosensors: more sensitive vs less sensitive

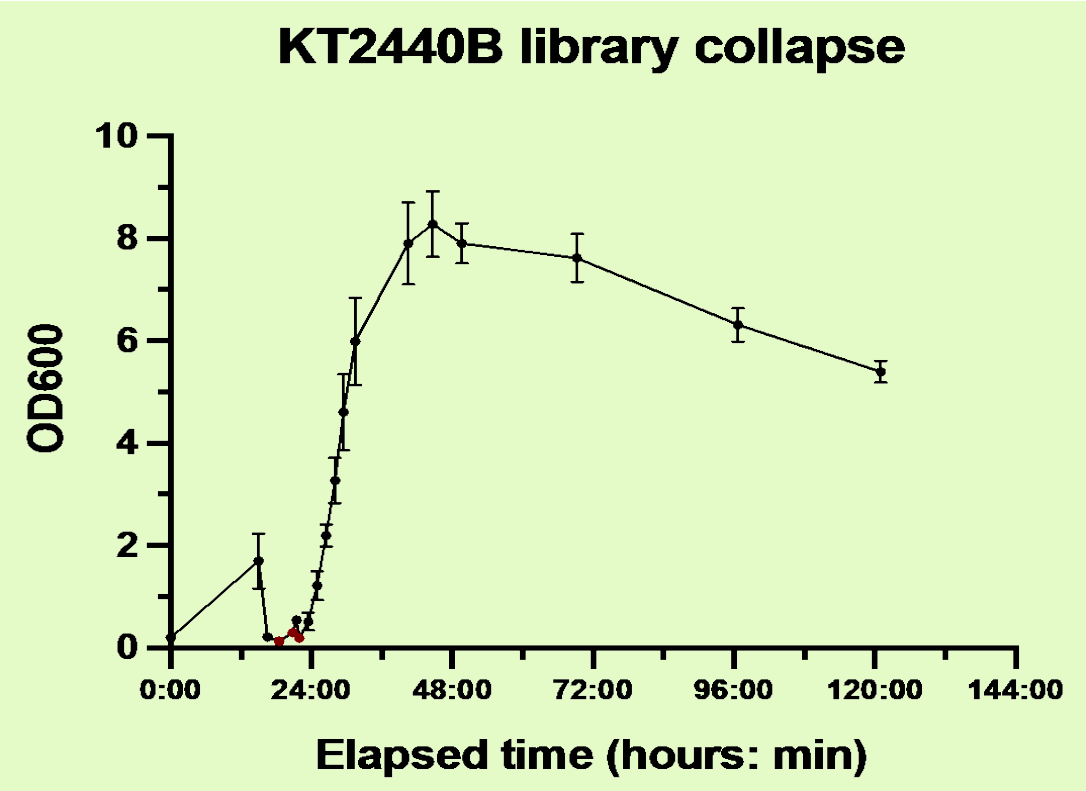


Monoclonal strain growth curves





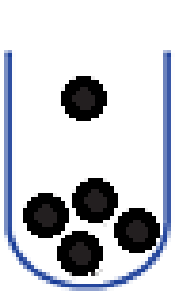
KT2440 Library collapsing experiment



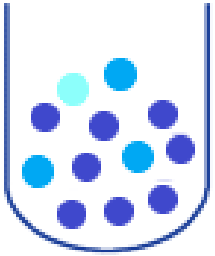
Purpose: to track library changes throughout growth.

- M9, 50mM glucose
- Performed in triplicate
- Samples taken for each time point shown.
- *MiSeq analysis pending protocol revisions.*

General experimental workflow plan

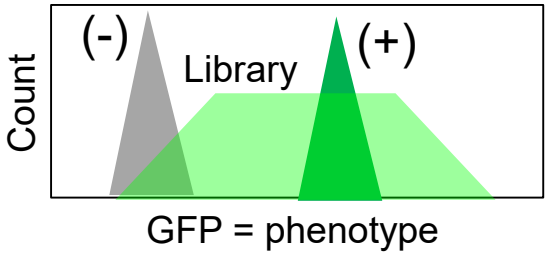


Transform with plasmid library

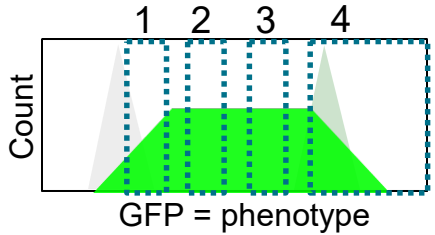


Generate population of varied phenotypes

Tie phenotype to fluorescent readout

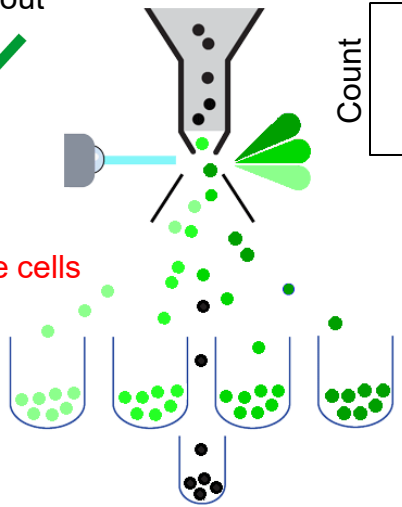


(-) WT KT2440
(+) Parent strain



Define populations to be collected by "gating"

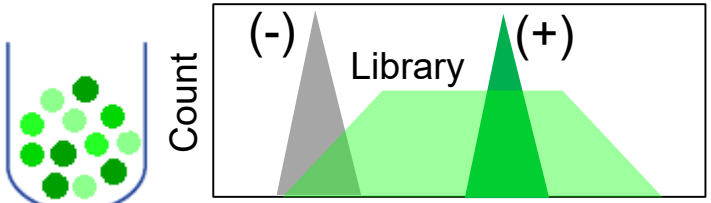
Cell sorting to isolate cells



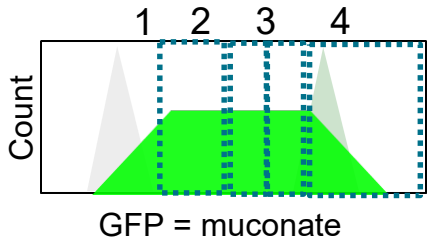
- Plasmid/amplicon sequencing of each sorted population to look for enriched sequences that can be correlated to each phenotype
- HTP LC-MS (SIM) for titers

Cell sorting plan, added detail

- Sort all 3 triplicates 4 ways in parallel = 15 sorts & 60 samples - 2 full days of sorts
- ~1 x 10^6 cells collected per sort

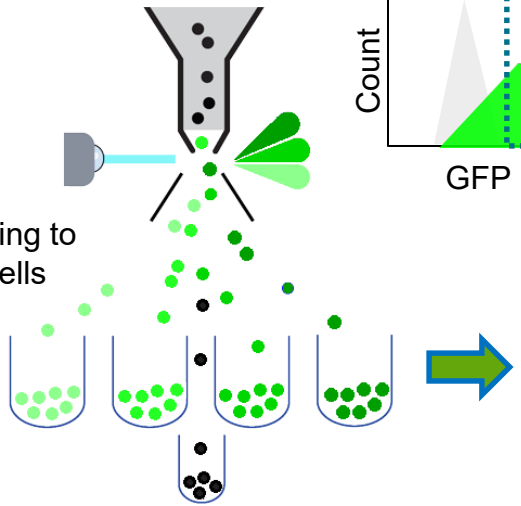


(-) WT KT2440
(+) Parent strain



Select 4 gates
Keep simple 1st try through Learn ("good vs bad")

Cell sorting to isolate cells

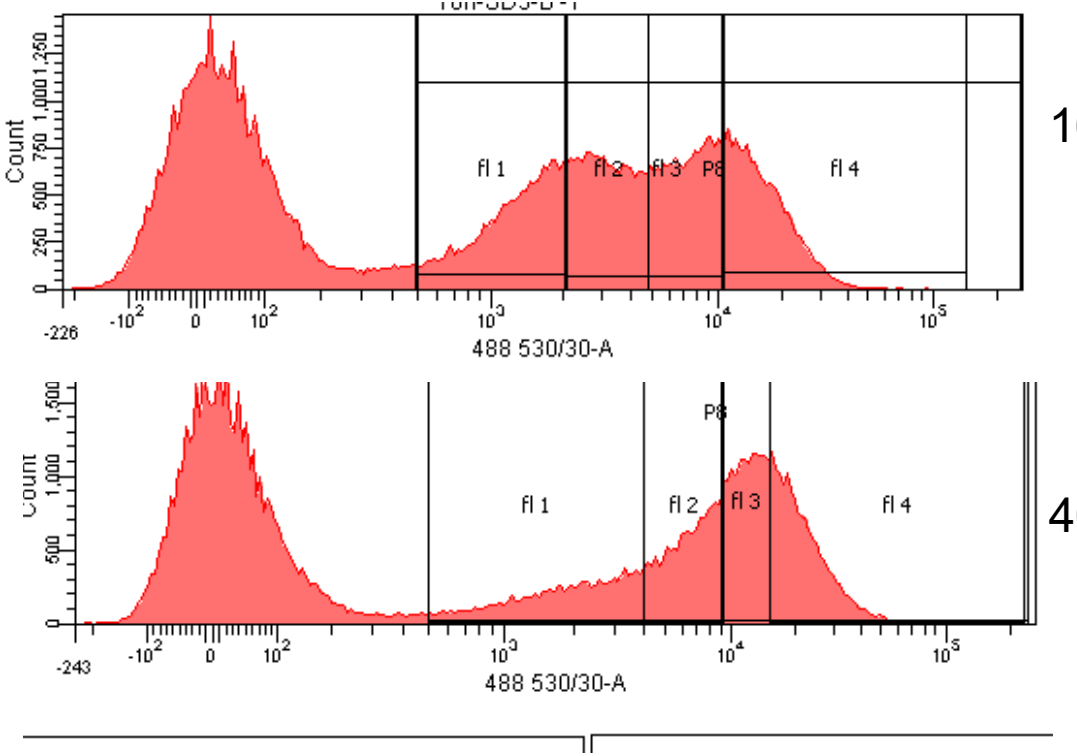


transrepressor length = 100 (200+ with terminator)			Actuals
population	6000	transrepressors -- 1 designed per gene	5500
desired sequencing coverage	300	x	182
need to collect	1.80E+06	cells	1.00E+06
FACS optimal event rate	8000	cells/sec	8000
# of parallel sorts	4		4
Fraction of population gated	50%	Fraction of population gated	64%
cells per sort (cells/m)	1000	cells/sec into each tube	1280
	1800	sec	781
	30	min of sorting	13

- Sequencing of sorted populations to look for enriched sequences that can be correlated to each phenotype
- PCR amplicon-based
- HTP LC-MS (SIM) of populations

U.S. DEPARTMENT OF ENERGY | Energy Efficiency & Renewable Energy
BIOENERGY TECHNOLOGIES OFFICE

FACS of trans repressor library SD3

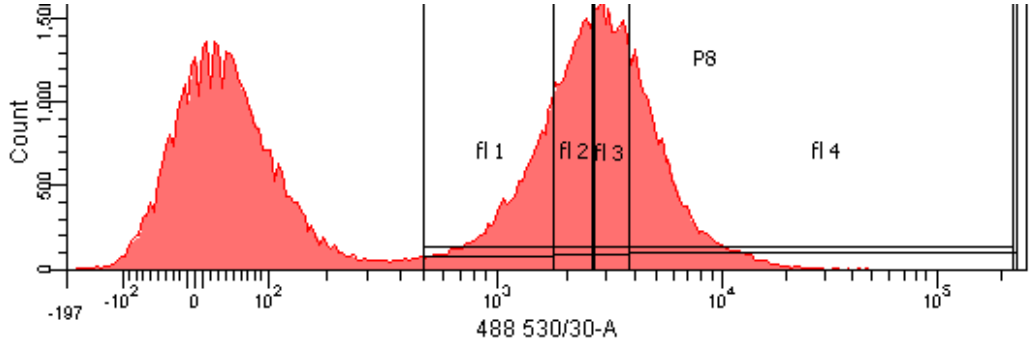


- Aria data with four gates
- Actual 1M cells collected per gate
- How collected (tubes)
- 55% of population collected

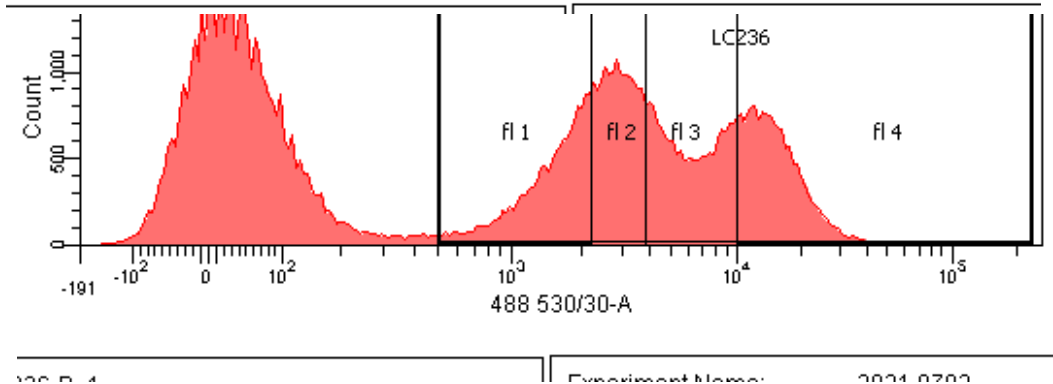
U.S. DEPARTMENT OF ENERGY | Energy Efficiency & Renewable Energy
BIOENERGY TECHNOLOGIES OFFICE

Agile BioFoundry

FACS of trans repressor library LC236



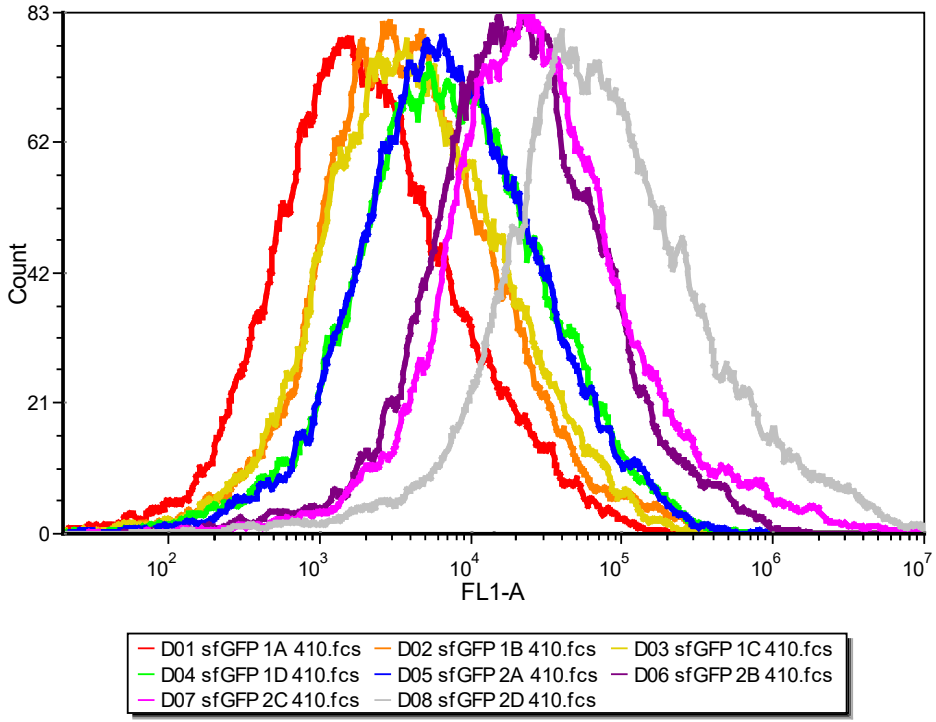
16h



40h

- Aria data with four gates
- Actual 1M cells collected per gate
- How collected (tubes)
- 57% of population collected

Example: analytical flow cytometry of sorted populations

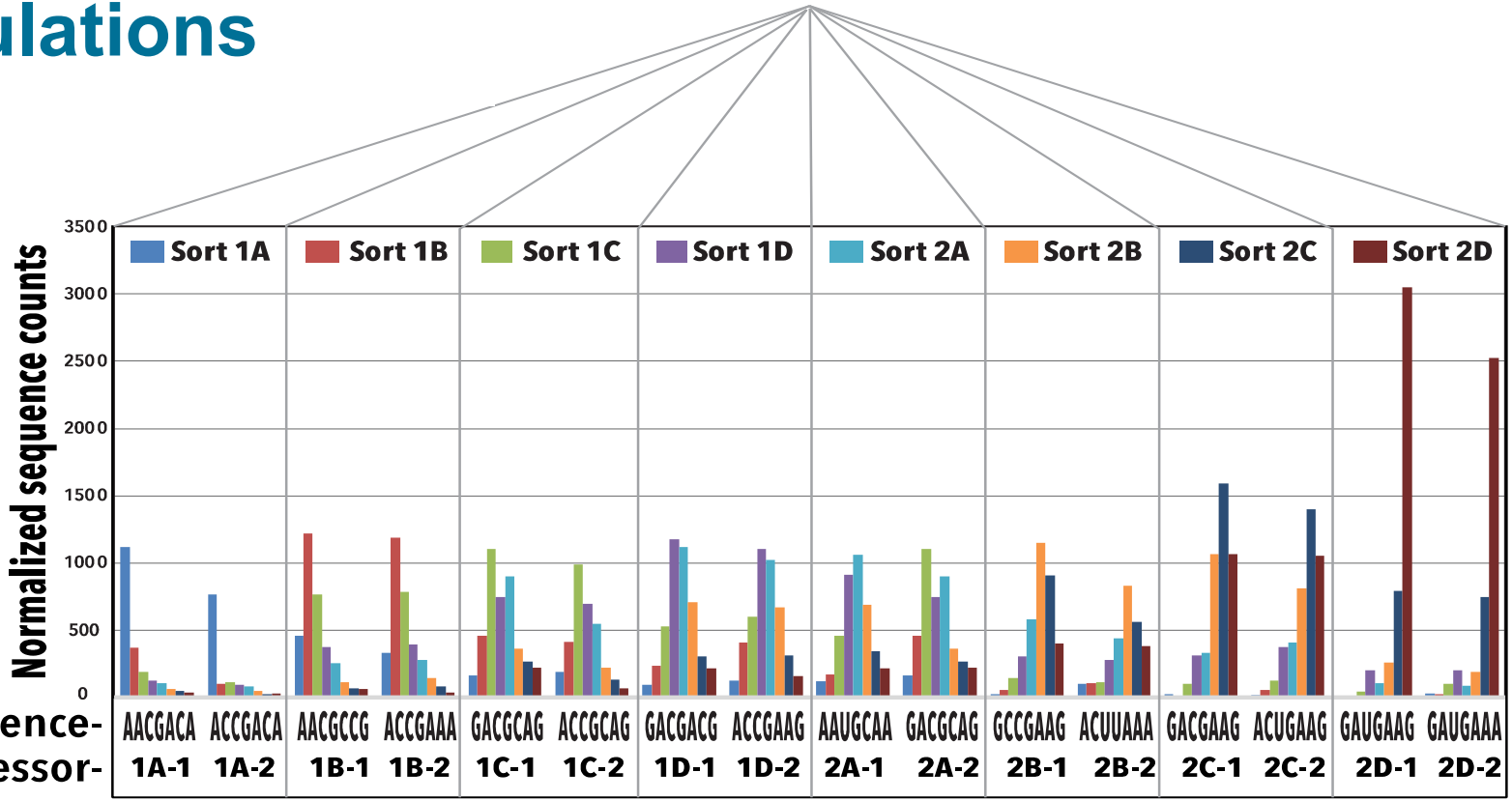


- Overnight re-growth and analysis shows sorted populations remain (mostly) distinctive
- Plasmid miniprep
- PCR barcode+ illumina adapters
- Sequencing

Example: sequence selection in sorted populations

- Sequences selected “by eye/hand”

Deep-sequence sorted pools



Comments from meeting

- LC224-B FACs – re-draw the gates!
- Sort NP162-B when library is completed (with and without gates)
- Amplicon sequencing vs plasmid sequencing
  - Go ahead with amplicon sequencing for the samples
  - Send pooled samples (SD3-B, LC224-B, LC236-B and LC238-B) and some separate samples (NP162 samples)
  - Send PCR cycles (10-40 cycles) for sequencing (determine PCR bias)
  - LBL can do the bulk 100+ sequencing samples, Argonne can do a few sequencing samples
  - Work with Josh McCaully, not Garima (left LBL)
  - Phil will check with Gyorgy about cycle number (usually less than 15 cycles, no PCR reach exponential phase yet)
  - Nathan thinks they will sequence whole 450bp
- Productivity vs yield muconate (perform a time-course with FI, muconate yield and OD of the selected strains)
- Share on a Google Drive (Jeremy will send a place to port if over) - old sequencing data to either Jeremy Zucker and Hector
- Jeremy will set up a meeting with Craig Bakker early Oct for stats analysis on old data)
- Send samples mid-October to LBL
- If we need a re-sort, then we can quickly get through it as we have a setup pipeline (milestone)
- Sort with gaps between the gates (maybe do NP162-B)
- Hector needs to get together Ramesh to get the background for the experiment
- Next meeting will revolved around Hector's understandings and aims of the project (maybe combination of genes in the future, maybe knockout strains and in combination)